Interpretazione "fisica" dei tassi di transizione

Riprendiamo l'equazione di Chapman-Kolmogorov nel caso di una catena di Markov a tempo continuo omogenea:

$$\frac{dH(\tau)}{d\tau} = H(\tau)Q, \quad \text{con } H(0) = I \Rightarrow \begin{cases} p_{i,i}(0) = 1 \\ p_{i,r}(0) = 0 \quad \forall r \neq i \end{cases}$$

Per l'elemento di posto (i,j), con $j \neq i$:

$$\frac{dp_{i,j}(\tau)}{d\tau} = \sum_{r \in \mathcal{X}} p_{i,r}(\tau)\, q_{r,j}$$

prodotto della i-esima riga di (τ) per la j-esima colonna di Q

$$= \underbrace{p_{i,i}(\tau)}_{\substack{\parallel \\ 1 \text{ per } \tau=0}} q_{i,j} + \sum_{r \neq i} \underbrace{p_{i,r}(\tau)}_{\substack{\parallel \\ 0 \text{ per } \tau=0}} q_{r,j}$$

Dunque:

$$\boxed{\left.\frac{dp_{i,j}(\tau)}{d\tau}\right|_{\tau=0} = q_{i,j}}$$

interpretazione

$q_{i,j}$ è il tasso istantaneo con cui (in senso probabilistico) può avvenire una transizione da i a j, $j \neq i$.

Consideriamo ora $\tau > 0$ sufficientemente piccolo cosicché nell'intervallo $(t, t+\tau]$ può avvenire una sola transizione di stato

(si ricordi che abbiamo assunto che due transizioni di stato non possono accadere nello stesso istante di tempo).

Supponiamo di entrare nello stato i all'istante t, e chiamiamo $V_{i,j}$ la durata di vita dell'evento $e_{i,j}$, per ogni $e_{i,j} \in \Gamma(i)$. Ricordiamo che $V_{i,j}$ segue una distribuzione esponenziale con parametro $\lambda_{i,j}$.

Segue che:

$$p_{i,j}(\tau) = P\big(X(t+\tau) = j \mid X(t) = i\big) =$$

sotto la nostra assunzione sull'ampiezza τ dell'intervallo

$$= P\big(V_{i,j} \leq \tau,\ V_{i,r} > V_{i,j}\ \ \forall e_{i,r} \neq e_{i,j}\big)$$

$$= \ldots$$

$$= \frac{\lambda_{i,j}}{\Lambda(i)}\left(1 - e^{-\Lambda(i)\tau}\right)$$

Questo implica che:

$$\frac{d\,p_{i,j}(\tau)}{d\tau} = \lambda_{i,j}\, e^{-\Lambda(i)\tau}$$

e quindi

$$\boxed{\left.\frac{d\,p_{i,j}(\tau)}{d\tau}\right|_{\tau=0} = \lambda_{i,j}}$$

Abbiamo cioè dimostrato che $\boxed{q_{i,j} = \lambda_{i,j}}$.

<u>interpretazione</u>

$q_{i,j}$ è il <u>tasso</u> della distribuzione esponenziale seguita dalle durate di vita dell'evento $e_{i,j}$, $j \neq i$

Possiamo ripetere gli stessi calcoli per $j=i$:

$$\frac{d\,p_{i,i}(\tau)}{d\tau} = \underbrace{p_{i,i}(\tau)}_{\substack{\shortparallel \\ 1 \text{ per } \tau=0}} q_{i,i} + \sum_{r\neq i} \underbrace{p_{i,r}(\tau)}_{\substack{\shortparallel \\ 0 \text{ per } \tau=0}} q_{r,i}$$

Dunque:

$$\left.\frac{d\,p_{i,i}(\tau)}{d\tau}\right|_{\tau=0} = q_{i,i}$$

Conviene riscrivere la precedente relazione come:

$$\boxed{\left.\frac{d}{d\tau}\left[1-p_{i,i}(\tau)\right]\right|_{\tau=0} = -q_{i,i}}$$

Si osservi che $1-p_{i,i}(\tau)$ e' la probabilita' (assumendo τ sufficientemente piccolo) che la catena abbandoni lo stato i.

<u>interpretazione</u>

> $-q_{i,i}$ e' il <u>tasso istantaneo</u> con cui (in senso probabilistico) la catena puo' abbandonare lo stato i.

=> $-q_{i,i}$ descrive la propensione della catena ad abbandonare lo stato i.

Inoltre, considerando τ sufficientemente piccolo, come sopra:

$$p_{i,i}(\tau) = P\big(X(t+\tau)=i \mid X(t)=i\big) =$$

sotto la nostra assunzione sull'ampiezza τ dell'intervallo

$$= P\big(V(i) > \tau\big) = e^{-\Lambda(i)\tau}$$

tempo di soggiorno nello stato i; si ricordi che $V(i)$ segue una distribuzione esponenziale con parametro $\Lambda(i)$.

Questo implica che:

$$\frac{d p_{i,i}(\tau)}{d\tau} = -\Lambda(i)\, e^{-\Lambda(i)\tau}$$

e quindi:

$$\boxed{\left.\frac{d p_{i,i}(\tau)}{d\tau}\right|_{\tau=0} = -\Lambda(i)}$$

Abbiamo cioe` dimostrato che $\boxed{q_{i,i} = -\Lambda(i)}$

Interpretazione

- $q_{i,i}$ rappresenta il tasso della distribuzione esponenziale seguita dal tempo di soggiorno nello stato i

OSSERVAZIONI

i) Dato che $\Lambda(i) > 0$, e abbiamo mostrato che $q_{i,i} = -\Lambda(i)$, segue che $q_{i,i} < 0$.

ii) Dato che $\lambda_{i,j} > 0$ per tutti gli eventi $e_{i,j} \in \Gamma(i)$, e abbiamo mostrato che $q_{i,j} = \lambda_{i,j}$, segue che:

- $q_{i,j} > 0$ se $e_{i,j} \in \Gamma(i)$
- $q_{i,j} = 0$ se $e_{i,j} \notin \Gamma(i)$

$j \neq i$

iii) Dato che $\Lambda(i) = \sum_{e_{i,j} \in \Gamma(i)} \lambda_{i,j}$, risulta che

$$-q_{i,i} = \sum_{j \neq i} q_{i,j}, \text{ e quindi } \boxed{\sum_j q_{i,j} = 0}$$

Concludiamo che, per una catena di Markov a tempo continuo omogenea, la <u>matrice Q dei tassi di transizione</u> è tale che:

i) tutti gli elementi lungo la diagonale principale sono negativi.

ii) tutti gli altri elementi sono non negativi.

iii) la somma lungo una qualsiasi riga di Q fa 0.

<u>CONSEGUENZA</u> - Se Q ha dimensioni finite, allora $\lambda=0$ è autovalore di Q.

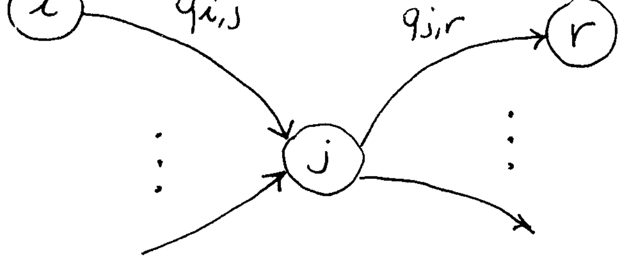
Infatti, si prenda il vettore $v=[1\ 1 \dots 1]^T$. Risulta:

$$Qv = \begin{bmatrix} \vdots \\ \sum_j q_{i,j} \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0$$

$\Rightarrow$ v è autovettore di Q relativo all'autovalore $\lambda=0$.

Rappresentazione grafica

Una catena di Markov a tempo continuo omogenea può essere rappresentata con un diagramma dei tassi di transizione dello stato (grafo) in cui gli archi sono etichettati con i tassi di transizione. Non si mette l'arco da i a j se $q_{i,j}=0$; non si mettono "riccioli".



transizioni da uno stato in se stesso non sono viste come eventi.

Probabilità di transizione (in un passo)

Supponiamo che le transizioni di stato (gli eventi della catena) avvengano in istanti aleatori $T_1 < T_2 < \ldots < T_k < \ldots$, e chiamiamo X_k lo stato dopo la k-esima transizione di stato (X_0 è lo stato iniziale).

Poniamo:

$$p_{i,j} = P(X_{k+1} = j \mid X_k = i) \quad \text{- probabilità di transizione in un passo -}$$

Abbiamo, per $j \neq i$:

$$p_{i,j} = P(X_{k+1} = j \mid X_k = i) = P(\text{evento } e_{i,j}) = \frac{\lambda_{i,j}}{\Lambda(i)} = \frac{q_{i,j}}{-q_{i,i}}$$

Mentre, per $j = i$:

$$p_{i,i} = P(X_{k+1} = i \mid X_k = i) = 0 \quad \text{perchè una transizione cambia lo stato.}$$

verifica:

$$p_{i,i} = 1 - P(X_{k+1} \neq i \mid X_k = i) = 1 - \sum_{j \neq i} P(X_{k+1} = j \mid X_k = i)$$

$$= 1 - \sum_{j \neq i} p_{i,j} = 1 - \sum_{j \neq i} \frac{q_{i,j}}{-q_{i,i}} = 1 + \frac{1}{q_{i,i}} \sum_{j \neq i} q_{i,j}$$

$$= 1 + \frac{1}{q_{i,i}} (-q_{i,i}) = 0$$

Per una catena di Markov a tempo continuo omogenea la matrice delle probabilità di transizione in un passo assume la forma:

$$P = [p_{i,j}]_{\substack{i=0,1,\ldots \\ j=0,1,\ldots}} = \begin{bmatrix} 0 & \frac{q_{0,1}}{-q_{0,0}} & \frac{q_{0,2}}{-q_{0,0}} & \cdots \\ \frac{q_{1,0}}{-q_{1,1}} & 0 & \frac{q_{1,2}}{-q_{1,1}} & \cdots \\ \frac{q_{2,0}}{-q_{2,2}} & \frac{q_{2,1}}{-q_{2,2}} & 0 & \ddots \\ \vdots & & & \end{bmatrix}$$

ATTENZIONE: A tempo discreto, con "un passo" si intende un istante temporale discreto. Dunque, in tal caso, gli elementi sulla diagonale di P (matrice di transizione a un passo) possono anche essere non nulli.

A tempo continuo, con "un passo" si intende un evento transizione di stato, e quindi gli elementi sulla diagonale di P sono tutti nulli.

Una catena di Markov a tempo continuo omogenea è completamente specificata da:

- $\mathcal{X}$: spazio degli stati
- Q : matrice dei tassi di transizione
- $\pi(0)$: vettore delle probabilità degli stati all'istante iniziale

In alternativa a Q, si può fornire la matrice P delle probababilità di transizione in un passo e, in aggiunta, i parametri delle distribuzioni esponenziali dei tempi di soggiorno in ciascuno stato (valori $-q_{i,i}$).

Probabilità degli stati e analisi del transitorio

Definiamo

$$\pi_i(t) = P(X(t) = i)$$ - probabilità dello stato i all'istante t -

e

$$\pi(t) = [\pi_0(t)\ \pi_1(t)\ \dots]$$ - vettore (riga) delle probabilità degli stati -

Osserviamo che:

$$\pi_j(t) = P(X(t)=j) = \sum_{i \in \mathcal{X}} \underbrace{P(X(t)=j \mid X(0)=i)}_{p_{i,j}(t)} \underbrace{P(X(0)=i)}_{\pi_i(0)}$$

$$= \sum_{i \in \mathcal{X}} \pi_i(0)\, p_{i,j}(t)$$

prodotto di $\pi(0)$ per la colonna j-esima di $H(t)$.

Si ricordi che $H(t) = \left[p_{i,j}(t) \right]_{\substack{i=0,1,\dots \\ j=0,1,\dots}}$

Dunque, in forma matriciale:

$$\pi(t) = \pi(0) H(t)$$

e ricordando che $H(t) = e^{Qt}$:

$$\boxed{\pi(t) = \pi(0)\, e^{Qt}}$$

Cioè, $\pi(t)$ è soluzione dell'equazione differenziale

$$\boxed{\dot{\pi}(t) = \pi(t) Q}$$

con condizione iniziale $\pi(0)$.

Riepilogo

Due equazioni differenziali <u>fondamentali</u> per catene di Markov a tempo continuo omogenee:

i) $\dfrac{dH(\tau)}{d\tau} = H(\tau) Q \quad \Rightarrow \quad H(\tau) = e^{Q\tau}$

ii) $\dot{\pi}(t) = \pi(t) Q \quad \Rightarrow \quad \pi(t) = \pi(0) e^{Q\tau}$

La soluzione di queste equazioni ci permette di svolgere <u>l'analisi del transitorio</u> per catene di Markov a tempo continuo omogenee.

ANALISI A REGIME

Dato uno stato i di una catena di Markov a tempo continuo omogenea, consideriamo il limite

$$\pi_i = \lim_{t\to\infty} \pi_i(t) \qquad \text{dove } \pi_i(t) = P(X(t)=i)$$

Se π_i esiste, si dice probabilità stazionaria (o a regime) dello stato i.

Se π_i esiste per ogni $i \in \mathcal{X}$, il vettore

$$\pi = [\pi_0 \; \pi_1 \; \pi_2 \; \dots]$$

si dice vettore delle probabilità stazionarie dello stato.

ATTENZIONE: la quantità che raggiunge regime è la probabilità, non lo stato, che è sempre una variabile aleatoria.

Ci poniamo le seguenti domande:

i) Fissato $i \in \mathcal{X}$, sotto quali condizioni π_i esiste?

ii) Se π_i esiste per ogni $i \in \mathcal{X}$, $\{\pi_i\}_{i=0}^{\infty}$ definisce una densità di probabilità discreta, cioè $\sum_{i=0}^{\infty} \pi_i = 1$?

NOTA: $0 \le \pi_i \le 1$ perchè $0 \le \pi_i(t) \le 1$ per ogni t.

Solo in questo caso i valori $\{\pi_i\}$ sono interpretabili come la densità di probabilità stazionaria dello stato della catena.

iii) Come valutare π_i?

OSSERVAZIONE: L'analisi del comportamento a regime per catene di Markov a tempo continuo omogenee è molto simile a quella svolta nel caso a tempo discreto. In particolare, i concetti di irriducibilità e ricorrenza sono ancora validi, ecc.

Enunciamo solo il risultato fondamentale:

TEOREMA - In una catena di Markov irriducibile in cui tutti gli stati sono ricorrenti positivi, esiste un unico vettore delle probabilità stazionarie dello stato, $\Pi = [\Pi_0 \; \Pi_1 \; \Pi_2 \; \dots]$. Esso è tale che $\Pi_i > 0$ e $\Pi_i = \lim_{t \to \infty} \pi(t)$ per ogni $i \in \mathcal{X}$, indipendentemente dal vettore delle probabilità dello stato all'istante iniziale, $\pi(0)$. Il vettore Π si può ottenere risolvendo il sistema di equazioni

$$\begin{cases} \Pi Q = 0 \\ \sum_{i \in \mathcal{X}} \Pi_i = 1 \end{cases}$$

OSSERVAZIONE - Si ricordi che, se la catena di Markov è finita, $\lambda = 0$ è autovalore della matrice dei tassi di transizione Q, e quindi il sistema

$$\Pi Q = 0$$

ha infinite soluzioni perché Q è singolare. Sotto le ipotesi del teorema, il vincolo $\sum_{i \in \mathcal{X}} \Pi_i = 1$ permette di selezionare l'unica soluzione che rappresenta una densità di probabilità discreta.

OSSERVAZIONE - Analogamente al caso a tempo discreto, è possibile mostrare che ogni catena di Markov a tempo continuo irriducibile e finita possiede un unico vettore delle probabilità stazionarie dello stato.